

Psychological  
Monographs

General and Applied

UNIVERSITY  
OF MICHIGAN

APR 7 1954

SCIENCE  
LIBRARY

A Prognostic Scale for  
Shock Therapy

By

Marvin J. Feldman

Marvin J. Feldman

Vol. 65  
No. 10

Edited by Herbert S. Conrad

Published by The American Psychological Association



# Psychological Monographs: General and Applied

## *Editor*

HERBERT S. CONRAD

Federal Security Agency  
Office of Education  
Washington 25, D.C.

## *Managing Editor*

LORRAINE BOUTHILET

## *Consulting Editors*

DONALD E. BAIER

FRANK A. BEACH

ROBERT G. BERNREUTER

WILLIAM A. BROWNELL

HAROLD E. BUETT

JERRY W. CARTER, JR.

CLYDE H. COOMBS

JOHN G. DARLEY

JOHN F. DASHIELL

EUGENIA HANFMANN

EDNA HEIDREDER

HAROLD E. JONES

DONALD A. MACKINNON

LORRIN A. RIGGS

CARL R. ROGERS

SAUL ROSENZWEIG

ROSS STAGNER

PERCIVAL M. SYMONDS

JOSEPH TIFFIN

LEDYARD R. TUCKER

JOSEPH ZUBIN

MANUSCRIPTS should be sent to the Editor. For suggestions and directions regarding the preparation of manuscripts, consult the following article: CONRAD, H. S., Preparation of manuscripts for publication as monographs. *J. Psychol.*, 1948, 26, 447-459.

Because of lack of space, the *Psychological Monographs* can print only the original or advanced contribution of the author. *Background and bibliographic materials must, in general, be totally excluded*, or kept to an irreducible minimum. Statistical tables should be used to present only the most important of the statistical data or evidence.

CORRESPONDENCE CONCERNING BUSINESS MATTERS (such as subscriptions and sales, change of address, author's fees, etc.) should be addressed to the American Psychological Association, 1515 Massachusetts Ave., N.W., Washington 5, D.C.

VOLUME 65  
NUMBER 10

WHOLE No. 327  
1951

# Psychological Monographs:

## General and Applied

Combining the *Applied Psychology Monographs* and the *Archives of Psychology*  
with the *Psychological Monographs*

HERBERT S. CONRAD, *Editor*

---

## A Prognostic Scale for Shock Therapy

By

MARVIN J. FELDMAN  
*University of Buffalo*

This paper is a revision of the dissertation submitted to the  
Department of Psychology of the University of California  
in partial fulfillment of the requirements for the  
degree of Doctor of Philosophy, January 1949

*Accepted for publication, November 10, 1950*

Price \$1.00

*Published by*

THE AMERICAN PSYCHOLOGICAL ASSOCIATION  
1515 MASSACHUSETTS AVE., N.W., WASHINGTON 5, D.C.

Psychological Monographs  
General and Applied

COPYRIGHT, 1952, BY THE  
AMERICAN PSYCHOLOGICAL ASSOCIATION

## TABLE OF CONTENTS

I. INTRODUCTION .....	1
II. THE DERIVATION OF THE <i>Ps</i> SCALE .....	3
A. The Plan of Investigation .....	3
B. The Criterion Groups .....	4
C. Item Analysis .....	7
D. The <i>Ps</i> Scale Items .....	7
III. APPLICATION OF <i>Ps</i> TO VARIOUS GROUPS .....	9
A. Application of <i>Ps</i> to the Criterion Groups and to New Test Groups .....	9
B. Application of <i>Ps</i> to a Group of Normals .....	12
C. Application of <i>Ps</i> to a Group of Psychoneurotics .....	12
D. Application of <i>Ps</i> to Postshock MMPI Records of the Test Groups .....	14
IV. ANALYSIS OF THE <i>Ps</i> SCALE .....	15
A. Reliability of the <i>Ps</i> Scale .....	15
B. Item Overlap and Intercorrelations .....	15
C. Relation of <i>Ps</i> to Type of Shock .....	17
D. Relation of <i>Ps</i> to Diagnosis .....	18
V. DISCUSSION AND INTERPRETATION OF <i>Ps</i> .....	20
VI. THE LOW-PROFILE GROUP .....	23
A. Scores on <i>Ps</i> .....	23
B. Item Analysis .....	23
VII. SUMMARY AND CONCLUSIONS .....	25
APPENDIX .....	26
REFERENCES .....	27





### ACKNOWLEDGEMENTS

**T**HE WRITER wishes to acknowledge his indebtedness to those persons who had a part in making this study possible. In particular he is grateful to Drs. Harold D. Carter, Donald W. MacKinnon, and Alexander P. Simon for their critical reading of the text, and above all to Dr. Robert E. Harris for his supervision and many helpful suggestions.

MARVIN J. FELDMAN

THE HISTORY OF THE  
LORDS OF THE MANOR OF  
ST. MARTIN'S, PARISH OF  
ST. MARTIN, IN THE  
COUNTY OF MIDDLESEX,  
FROM THE EARLIEST  
RECORDS TO THE PRESENT  
TIME.  
BY  
JAMES H. STUBBS, ESQ.  
OF THE BARR.

THE HISTORY OF THE LORDS OF THE MANOR OF ST. MARTIN'S, PARISH OF ST. MARTIN, IN THE COUNTY OF MIDDLESEX, FROM THE EARLIEST RECORDS TO THE PRESENT TIME. BY JAMES H. STUBBS, ESQ. OF THE BARR.



## CHAPTER I

### INTRODUCTION<sup>1</sup>

THE OBJECTIVE measurement of clinical status in psychopathology seems to be a necessary, although not a sufficient, condition to further progress in various types of therapy. The solution of many problems, such as the efficacy of a particular treatment, more refined classification of disease syndromes, the nature of the changes that take place during the course of treatment, etc., are contingent in part upon the development of more adequate measuring devices. Although the past decade has witnessed tremendous growth in the development of therapeutic techniques, it has produced few of the precise tests that are needed to predict and evaluate the results of the new methods of treatment.

One acute problem facing the present-day therapist is the choice of treatment for a particular patient. Should the patient be treated psychotherapeutically, e.g., by psychoanalysis, nondirective therapy, etc., or should a more drastic type of treatment be used, e.g., shock treatment, psychosurgery, etc.? In this situation some guiding principles based upon empirical evidence would be of material benefit. At the present time, however, the evidence is insufficient to evaluate the relative merits of the various treatments. On the basis of clinical experience we do have some notions about selecting patients for particular types of therapy. Some of the rules help to decide the suitability of a patient for one type of treatment, but alternative

choices are seldom specified. Thus a transference neurosis is generally assumed to be amenable to psychoanalytic treatment, but it is not known whether nondirective therapy would do as well. Then, too, although some of the rules undoubtedly are valid for groups of patients, they offer little direction in the individual case. For example, although drastic types of treatment are usually reserved for psychotic patients, it may be that a sizable minority of psychotics would benefit much more from some form of psychotherapy. But at present we have little more than clinical "hunches" about which psychotic patients would respond to psychotherapy. When the scope of comparison is narrowed to the various types of drastic treatment, there is at present no sound basis for deciding whether one type is more effective than the others for all cases or even for specific types of patients. Insulin is usually reserved for the most severe cases of schizophrenia, whereas electric shock is thought to work best in affective disorders; yet it is a moot point whether these notions would be substantiated in rigorous comparisons of matched groups. These few examples suffice to point out the need for objective instruments which could be used to test some of the conceptions held in current psychiatric practice.

The purpose of the present study is to derive empirically a scale which predicts the results of shock treatment. The proposed scale may have several uses:

1. *A practical use.* To obtain some means of selecting patients for whom

<sup>1</sup> This study was made possible by a research grant from the United States Public Health Service.

shock treatment would probably be successful.

2. *A theoretical use.* To facilitate a formulation concerning the differences in personality structure of those patients who respond well and those who respond poorly to shock treatment, in so far as these differences are reflected in changing verbal responses to test items.

3. *A heuristic use.* To furnish a ra-

tional basis for selecting patients for numerous possible comparative studies of therapy.

Thus, the present investigation is undertaken in the belief that an objective scale which predicts the results of shock treatment will not only have practical utility but will also be of theoretical interest.

## CHAPTER II

### THE DERIVATION OF THE *Ps* SCALE

THE PRESENT investigation attempted to derive a prognostic scale for shock therapy. The scale is labeled *Ps* in keeping with its intent, namely, prognosis for shock treatment. The Minnesota Multiphasic Personality Inventory (MMPI) furnished the initial pool of items for the *Ps* scale. The MMPI need not have been chosen as the pool of items for the new scale. In fact, a new pool of items could have been devised to reflect a more dynamic personality theory than do the items on the MMPI. Such a pool of items might have yielded even better discriminations than were obtained with the items actually used. The MMPI was used for two reasons:

1. Most important was a practical reason. A group of records was already available on a sizable number of patients treated with shock therapy. These records constituted a fairly representative sample of patients at the clinic.
2. Another reason was that a large body of knowledge had been accumulated on the results of the MMPI. In correlating *Ps* with these results, valuable insights might be gained into the nature of the new scale.

The MMPI has been described in detail elsewhere (4, 5, 7, 8, 9). Hence, only a brief recapitulation is presented in this paper. The inventory contains nine empirically derived personality variables, named as follows: *Hs* (hypochondriasis), *D* (depression), *Hy* (hysteria), *Pd*, (psychopathic deviate revised), *Mf* (masculinity-femininity), *Pa* (paranoia), *Pt* (psychasthenia), *Sc* (schizophrenia), and *Ma* (hypomania). Earlier forms of the *Hs* and *Pd*, scales were also used in the present study. In addition to the personality variables there are four measures of the test's internal validity. These scales detect subsets of individuals who for one reason or another do not sort the items with the same attitude as other persons.

The "?" score is simply the number of items answered "cannot say." Too large a "?" score impairs the validity of the entire test. The *L* (lie) scale is a set of more or less obvious items revealing tendencies of the respondent to put himself in too favorable a light. The *F* scale is composed of statistically infrequent responses and among other things reveals lack of understanding or uncooperativeness in taking the test. The most recently developed scale, *K*, detects more subtle test-taking attitudes of the respondent in his attempts to place himself in either a favorable or unfavorable light. Although the authors now recommend that certain of the personality scales be systematically corrected in terms of *K*, this procedure has not been followed in the present study. Some results in this investigation indicate that *K* has a limited application and would be inappropriate if used with the present groups.

#### A. THE PLAN OF INVESTIGATION

The following plan of investigation was employed in deriving the *Ps* scale:

1. Two criterion groups were selected, one of which responded well and the other poorly to shock treatment. The former is designated as the *recovered criterion group* and the latter as the *unimproved criterion group*.
2. An item analysis was performed on all 550 items of the MMPI. The items which were statistically significant in distinguishing between the criterion groups comprise the *Ps* scale.
3. The validity of the *Ps* scale was demonstrated by its application back to the original criterion groups and also to similar independent test groups, i.e., groups rated recovered, improved, and unimproved after shock treatment. Various other characteristics of *Ps* were elicited by its application to a normal sample, to a psychoneurotic group treated with brief psychotherapy, and to the post-shock records of the test groups.

4. Odd-even reliability coefficients were determined for the scale.

5. The item overlap and the inter-correlations between *Ps* and the other scales of the MMPI were calculated.

6. The scores on *Ps* were analyzed according to the type of shock treatment.

7. The scores on *Ps* were analyzed according to the diagnosis.

8. On the basis of item content and other findings, some theoretical formulations were offered concerning the differences in personality structure of those patients who responded well and those who responded poorly to treatment.

9. A separate item analysis was performed on a small group whose MMPI profiles did not deviate sufficiently in the direction of abnormality. The rationale for this procedure is explained on both clinical and statistical grounds.

#### B. THE CRITERION GROUPS

It is of paramount importance that criterion groups be selected with the greatest of care, for shoddy selection procedures can introduce unreliability into the scale. The cases selected should be as "pure" as possible in terms of the variable under consideration. If excessive doubt exists about a case, it should not be included in the criterion groups. With test groups, less rigorous procedures are in order. If the scale predicts accurately with loosely selected groups, it is an excellent indication of the test's validity.

The criterion groups were composed of two extreme groups, one of which responded very well and the other poorly to shock treatment. At the Langley Porter Clinic four ratings describe status at the termination of treatment: *recovered*, *much improved*, *improved*, and *unimproved*. Criteria for these ratings are as follows: A patient is regarded as *recov-*

*ered* if he returns to his premorbid status, is making a good social and occupational adjustment, and expresses good insight. He is regarded as *much improved* if he returns to a good social and occupational adjustment but shows residuals of his illness in such symptoms as emotional dullness, anxiety, etc. *Improved* indicates that there is some lessening of the intensity of the symptoms, but the fundamental process is unchanged. *Unimproved* means that the patient is essentially unchanged. A large majority of these ratings were made by a single, highly-skilled psychiatrist.

Ratings of *recovery*, *social recovery*, and *essentially unimproved* designate psychiatric status in subsequent follow-up studies done at three-, six-, and twelve-month intervals after the patient left the clinic. In most cases the patient was seen for a personal interview by a psychiatrist, but when this was not possible, the patient's condition was evaluated in a letter sent by the patient himself or by his relatives.

Various members of the staff made these ratings. Because the *much improved* category is rarely used, all patients were reclassified into three groups for the purposes of this study. To be placed in the *recovered criterion group*, a patient had to be rated recovered or much improved at the termination of shock treatment and had to maintain a recovery rating for one year after leaving the clinic. To be placed in the *unimproved criterion group*, a patient had to rate improved or unimproved at the termination of treatment and had to maintain an unimproved condition for one year after leaving the clinic. This leaves a middle group, an *improved group*. To be placed in the *improved group*, a patient had to be rated much



improved or improved at the termination of treatment and had to maintain a social recovery for one year after leaving the clinic. This latter group furnished one of the test groups and will be dealt with in a later section. Using follow-up studies to help classify patients insures greater accuracy by correcting the status of patients whose gains are transient. These patients are misclassified if status at the termination of treatment is the sole consideration.

A few patients could not be classified into any of the three groups because their psychiatric status was markedly unstable in the follow-up period. Such patients might furnish interesting material for study in themselves, but their number was too small for that purpose in the present investigation; hence, their records were discarded. Another small group of records was borderline, i.e., a record rated much improved at the termination of treatment and receiving both recovery and social recovery ratings in the follow-up period. In such cases the writer classified the record with the aid of the patient's chart, giving most weight to status at the termination of treatment. If there was too much doubt about a case, it was classified but set aside for use in the test groups.

No attempt was made to equate the criterion groups on the basis of various personal factors, for such requirements would seriously reduce the size of the groups. Also, the criterion groups were neither separated nor matched on the basis of the usual Kraepelinian nosology. Nor was the study restricted to one particular type of shock treatment; patients treated either with insulin or by electrical methods were included. The rationale for these procedures will be set forth in a later section.

One restriction in selecting patients was that the MMPI profile had to deviate sufficiently from normality. On the basis of both clinical and statistical experience, the writer felt that the small subgroup whose profiles did not deviate sufficiently from the norm should be analyzed separately. A record was included in the main criterion groups if it met either of the following requirements:

1. A *T*-score  $\geq 70$  on one or more of six scales, *Hs*, *D*, *Hy*, *Pa*, *Pt*, or *Sc*.

2. A *T*-score  $\geq 70$  on both *Pd*, and *Ma* even though all other scales were within normal limits (*Mf* being excluded from consideration).

A record was excluded if either *Pd*, or *Ma* was the only scale to reach a *T*-score  $\geq 70$ . On the basis of these requirements about 20 per cent of the records were set aside for special study.

A few cases were eliminated because of the following considerations of internal validity:

1. In order to rule out carelessness or lack of understanding on the part of the respondent, records were discarded if the raw score on *F* was above 21. Although this score is higher than the critical score set by the authors, experience with clinical groups shows that *F* legitimately attains higher magnitudes in psychiatric groups than in the general population (12).

2. Records with a "?" score in excess of 100 raw score points were eliminated because these records distort the item counts if included and, of course, are of questionable validity.

No restriction was placed upon the magnitude of either the *L* or *K* scores. In view of the fact that profiles must deviate significantly from normal to be included, the error introduced from high scores on these scales is immaterial.

With these restrictions in mind, a perusal of the clinic's files since 1943 yielded 184 MMPI tests administered from one week to one day before shock treatment was instituted. Of these, 42 form the recovered criterion group, and a like number the unimproved criterion group. Some personal data are presented to show that this sample is not unusual in any respect. The recovered criterion group range in age from 15 to 65, with a mean of 37.5 years, whereas the unimproved criterion group range in age from 15 to 59 with a mean of 32 years. Any advantage in age resides with the unimproved group. The unimproved group is almost equally divided in sex, 22 males and 20 females, while the recovered group contains a preponderance of 29 females to 13 males. It is difficult to assess the importance of this sex difference. An insignificant correlation of .11 between scores on *Ps* and the *Mf* scale of the MMPI furnishes presumptive evidence that the masculinity-femininity score of the respondent is not too important a factor in prognosis. Scores on an abbreviated form of the Wechsler-Bellevue Adult Intelligence Scale were available for part of each group. The four subtests used were Similarities, Comprehension, Block Design, and Digit Symbol. The average IQ of 33 members of the recovered group is 109.1, and the average IQ of 26 of the unimproved group is 103.1. The tests were administered before treatment. These results are similar to findings with other patients treated at the clinic (2).

Both groups contained a few patients who had received a previous course of shock treatment—4 in the recovered group and 6 in the unimproved group. The social consequences of unimprovement are quite evident in the fact that 26 of the 42 were subsequently commit-

ted to institutions. Only 9 of the recovered group had a similar fate. It should be noted that recovered patients were committed at least one year subsequent to the termination of treatment and usually enjoyed a recovered status for two to four years before relapsing.

Both groups received equally intense courses of treatment. The recovered group had an average of 12 electrically induced convulsions and 50 insulin comas, whereas the unimproved group had an average of 11 electrically induced con-

TABLE 1  
DIAGNOSES FOR THE TWO  
CRITERION GROUPS

Diagnosis	Criterion Group	
	Recovered	Unimproved
Affective disorders	29	9
Schizophrenia	11	22
Psychoneurosis	2	11
Total	42	42

vulsions and 42 insulin comas. A somewhat larger number of unimproved patients received insulin treatment (15 compared to only 7 in the recovered group).

The biggest differences between the two groups were in terms of diagnosis. Table 1 gives the distributions according to diagnosis. The discrepancies were probably not as marked as the results would tend to indicate, for to some degree diagnosis may be influenced by the patient's prognosis. This is especially true of patients with mixed symptoms; it becomes extremely difficult to decide whether the patient is manic-depressive or schizophrenic. It cannot be decided on an a priori basis whether or not these discrepancies introduce a serious methodological error. In a later section some empirical evidence will be presented to shed light on this problem.



## C. ITEM ANALYSIS

The methodology involved in the item analysis was as follows:

On the MMPI a "plus" response is defined as the direction in which an item is answered by the minority of the normal standardization group. Not all the items on the various scales of the MMPI are scored in the "plus" direction, however, because abnormal groups answer some items in the same way as do the majority of the normal standardization group, though to a statistically significant higher degree. These items are referred to as "zero" items. The first step was to sum the "plus" responses for each of the 550 items on the MMPI for both criterion groups. These sums were then converted to proportions. Thus, the percentage of "plus" responses to each item was obtained for each criterion group. The difference in proportion for each item was then analyzed to cull out items differentiating the groups at the 2 per cent level of confidence. The routine work was facilitated by using tables prepared by Edgerton and Paterson (1).

As already mentioned, the criterion groups contained 42 cases each. The question arises as to the optimal size of criterion groups. Strong (13) progressively increased his groups from 50 to 250 cases and found sufficient gains in the validity of his scales to recommend the latter number. On the MMPI the criterion groups have usually numbered around 40 to 50 cases. The argument advanced in support of the use of these smaller samples is that personality structure in diagnostic groups is more homogeneous than are interests in professional groups; hence, stable item counts can be achieved with smaller samples. It is a moot point whether or not further increases in the size of the present samples would have added appreciably to the validity of the scale. Exigencies of a finite sample precluded a definitive study of this point.

The item analysis of the recovered and unimproved criterion groups produced 52 items differentiating at the 2 per cent level of significance. Of the 52 items, 44

were significant at less than the 1 per cent level. These items comprise the scale labeled "Ps" (an abbreviation for "prognostic for shock"). The 52 items on the Ps scale are presented below. The letter *T* or *F* after each item indicates the direction in which the unimproved criterion group tends to respond in contrast to the recovered criterion group; hence, high scores on *Ps* indicate a poor prognosis.

D. THE *Ps* SCALE ITEMS

These items are listed according to their numbers as given in the back of the revised manual for the MMPI, 1943 (6).

- A-9. I think I feel more intensely than most people do. (T) (27, 40, 71)<sup>1</sup>
- A-50. I hardly ever feel pain in the back of my neck. (T) (21, 50, 28)
- B-25. I have never had any breaking out on the skin that has worried me. (F) (26, 24, 50)
- B-36. I dream frequently about things that are best kept to myself. (T) (24, 19, 50)
- B-48. I believe that my home life is as pleasant as that of most people I know. (F) (5, 19, 48)
- C-6. At times I have very much wanted to leave home. (T) (20, 31, 64)
- C-9. I have reasons for feeling jealous of one or more members of my family. (T) (9, 5, 24)
- C-10. Once in a while I feel hate towards members of my family whom I usually love. (T) (25, 26, 55)
- C-19. I am apt to pass up something that I want because others feel that I am not going about it in the right way. (T) (37, 24, 57)
- C-23. It bothers me to have someone watch me at work even though I know that I can do it well. (T) (49, 50, 73)
- C-24. I have often had to take orders from someone who did not know as much as I did. (T) (41, 17, 50)
- C-33. It takes a lot of argument to convince most people of the truth. (T) (59, 33, 62)
- C-35. I liked school. (F) (14, 14, 35)
- C-42. I like to read newspaper articles on crime. (T) (55, 21, 50)

<sup>1</sup> The numbers in parentheses after each item refer to the percentage of "plus" responses given to the item by the following groups, respectively: (a) the general MMPI normal group, (b) the recovered criterion group, and (c) the unimproved criterion group. Items J-3 and J-37 show the data for only the latter two groups.

- C-52. Many of my dreams have been about sex matters. (T) (8, 17, 45)
- D-1. I never attend a sexy show if I can avoid it. (F) (60, 38, 62)
- D-2. I like to talk about sex. (T) (74, 10, 38)
- D-45. People generally demand more respect for their own rights than they are willing to allow for others. (F) (76, 38, 12)
- D-47. Most people make friends because friends are likely to be useful to them. (T) (47, 21, 42)
- D-52. I think most people would lie to get ahead. (T) (43, 14, 42)
- E-2. There are certain people that I dislike so much that I am inwardly pleased when they are catching it for something they have done. (T) (25, 19, 45)
- E-11. At times I have been so entertained by the cleverness of a crook that I have hoped that he would get away with it. (T) (9, 14, 38)
- E-17. I have never been in trouble because of my sex behavior. (F) (15, 10, 35)
- E-48. I never worry about my looks. (F) (49, 35, 12)
- E-50. I do many things that I regret afterwards. (T) (38, 48, 76)
- E-52. People often disappoint me. (T) (56, 40, 71)
- F-2. In a group of people I would not be embarrassed to be called upon to start a discussion or give an opinion about something I know well. (F) (39, 40, 69)
- F-5. I wish I were not so shy. (T) (38, 48, 71)
- F-6. I am not unusually self-conscious. (F) (31, 57, 80)
- F-10. I am more sensitive than other people. (T) (24, 55, 78)
- F-13. I have sometimes stayed away from another person because I feared doing or saying something that I might regret afterward. (T) (41, 31, 66)
- F-16. I have a daydream life about which I do not tell other people. (T) (22, 14, 45)
- F-52. My feelings are not easily hurt. (F) (46, 59, 24)
- G-6. I have certainly had more than my share of things to worry about. (T) (37, 38, 66)
- G-20. Once a week or oftener I become very excited. (T) (11, 24, 52)
- G-22. Sometimes I become so excited that I find it hard to get to sleep. (T) (49, 48, 78)
- G-43. Bad words, often terrible words come into my mind and I cannot get rid of them. (T) (12, 21, 42)
- H-2. I have often felt that strangers were looking at me critically. (T) (34, 40, 71)
- H-9. I tend to be on my guard with people who tend to be somewhat more friendly than I had expected. (T) (68, 57, 33)
- H-10. I commonly wonder what reason another person may have for doing something nice for me. (T) (46, 33, 57)
- H-16. People say insulting things about me. (T) (9, 14, 38)
- H-52. I have no dread of going into a room by myself where people have already gathered and are talking. (F) (27, 35, 69)
- I-8. No one cares much what happens to you. (T) (12, 12, 35)
- I-9. Any man who is able and willing to work hard has a good chance of succeeding. (F) (9, 5, 24)
- I-14. I frequently ask people for advice. (T) (72, 42, 17)
- I-20. People can pretty easily change me even though I thought my mind was made up on a subject. (T) (27, 28, 52)
- I-27. I find it hard to keep my mind on a task or a job. (T) (10, 62, 14)
- I-35. Life is a strain for me most of the time. (T) (17, 57, 83)
- J-3. Usually I would prefer to work with a woman. (F) (-, 38, 17)
- J-37. I used to like hop scotch. (F) (-, 76, 50)
- J-41. Once in a while I think about things too bad to talk about. (T) (58, 64, 38)
- J-47. I do not like everyone I know. (T) (16, 33, 10)

Most of the items are answered in the "plus" direction by the unimproved group. Only nine of the items are scored as "zero" items, namely, A-50, D-45, E-48, H-9, I-14, J-3, J-37, J-41, and J-47. Hence, to accumulate a poor prognosis score, the respondent must give many statistically deviant answers.

### CHAPTER III

#### APPLICATION OF $P_s$ TO VARIOUS GROUPS

##### A. APPLICATION OF $P_s$ TO THE CRITERION GROUPS AND TO NEW TEST GROUPS

THE FIRST step was to apply  $P_s$  back to the original criterion groups. Table 2 shows that  $P_s$  discriminates the criterion groups quite successfully. This is a necessary but not a sufficient condition for establishing validity. The scale must also be able to discriminate new test groups. The test groups used in this study are as similar

doubtful cases, that is, the cases which did not have an unequivocal rating of outcome, were also included. In spite of the more lenient standards approximately 90 per cent of the test cases had been the subjects of follow-up study for at least a six-month period. The test groups include an improved group as well as recovered and unimproved groups. In all, there are 40 recovered, 40

TABLE 2  
COMPARISON OF MEANS OF CRITERION GROUPS ON  $P_s$

Group	N	Mean	SD	Diff	CR	P
Recovered	42	17.0	6.3	13.9	9.65	.0001
Unimproved	42	30.9	6.7			

as possible to the original criterion groups, but, as mentioned previously, the criteria for inclusion of a case were not quite so stringent in the test groups. The restriction on the time limit in the follow-up period was relaxed so that recently treated cases and cases with incomplete follow-up data were included. The few

improved, and 20 unimproved test cases. In age, sex, and intelligence, the groups are similar to the criterion groups (see Appendix). There are, however, fewer manic-depressives in the unimproved test group. Fewer of the unimproved test group were subsequently hospitalized, a fact which undoubtedly reflects in part

TABLE 3  
COMPARISON OF MEANS OF TEST GROUPS ON THE  $P_s$  SCALE

Groups	N	Mean	SD	Diff	CR	P
Recovered vs. unimproved						
Recovered	40	17.6	6.1	12.0	8.33	.0001
Unimproved	20	29.6	4.9			
Recovered vs. improved						
Recovered	40	17.6	6.1	7.1	5.59	.0001
Improved	40	24.7	5.3			
Improved vs. unimproved						
Improved	40	24.7	5.3	4.9	3.50	.0002
Unimproved	20	29.6	4.9			

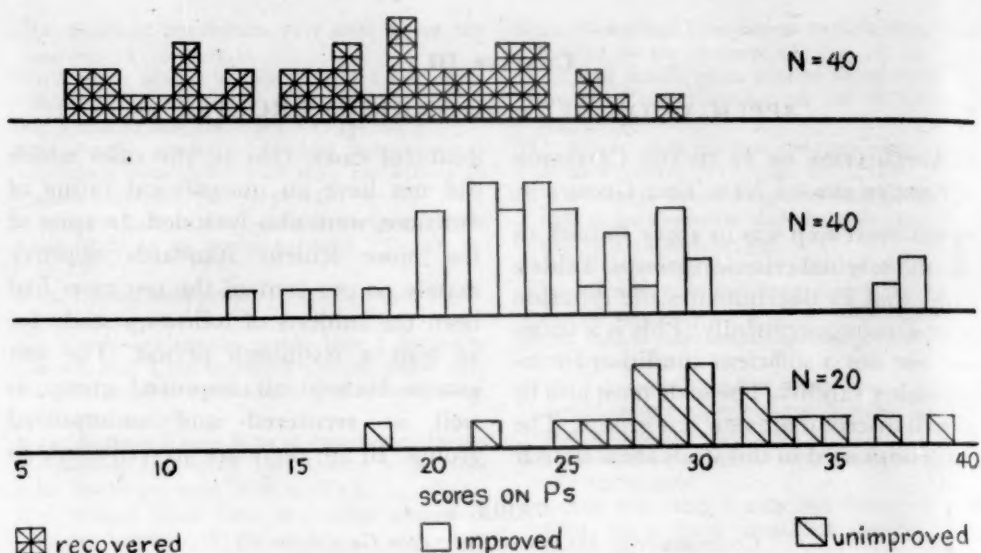


FIG. 1. Distributions of scores of test groups on the *Ps* scale

the shorter follow-up period. The differences in personal factors in the test groups contrasted to the criterion groups are all of such a nature as to introduce unreliability into the test scores of the test groups, hence making more exacting demands on *Ps*.

Table 3 shows that *Ps* works almost as well on the test groups as it does on the criterion groups. A graphic presentation of the discriminating power of *Ps* is portrayed in Figure 1. Although there are a few extreme "misses," the separations effected by *Ps* are about as clear-cut as can be expected from data rated with a reliability far short of perfection.

Another way of describing the predictive power of *Ps* is in terms of the overlap between the prognostic groups. Complete overlap is indicated, of course, when 50 per cent of one group  $\geq$  the median of another group. The amount of overlap is crucial in determining the practical clinical utility of a test. Yet a perusal of the psychological literature reveals that overlap is seldom reported. It

is almost as if most authors are unaware of the fact that a critical ratio is only the first step in establishing claims of validity for their tests or experiments. Table 4 indicates minimal overlap on *Ps* among the

TABLE 4  
PER CENT OVERLAP AMONG TEST GROUPS ON *Ps*

Group	Per Cent Overlap
Recovered $\geq$ Mdn of unimproved	0.0
Recovered $\geq$ Mdn of improved	18.0
Improved $\geq$ Mdn of unimproved	12.0
Unimproved $\leq$ Mdn of improved	10.0
Unimproved $\leq$ Mdn of recovered	5.0
Improved $\leq$ Mdn of recovered	10.0

various test groups. Naturally, the greatest amount of overlap is between the improved group and each of the two extreme groups. The writer does not intend to explain away the test "misses," but it would be interesting to follow the subsequent psychiatric status of the test misses to see if they conform to the expectation of the test. For example, three patients were rated improved clinically but re-



ceived high scores on *Ps*. Are these patients more likely to relapse than other patients rated improved clinically but scoring lower on *Ps*?

For practical clinical use, it is desirable to present the data in the form that can be most conveniently interpreted. On the MMPI, scales derived from clinical groups are scored on a large normal sample. The distribution of raw scores on the normal sample is then converted to a set

arising from the combined use of criterion and test cases should more than offset the slight loss of accuracy arising from this procedure.

A much larger sample would be desirable, of course, to furnish more stable percentages within the various bands of scores on *Ps*. Any given score can be easily interpreted from Table 5. If a patient receives a score of 29 on *Ps*, reading from the table we see that 2 per cent of

TABLE 5  
PERCENTAGE OF EACH TEST GROUP FALLING WITHIN  
PRESCRIBED ZONES ON THE *Ps* SCALE

Group	Raw Scores on <i>Ps</i>									
	5-8	9-12	13-16	17-20	21-24	25-28	29-32	33-36	37-40	41-44
Recovered	100%	100%	67%	57%	33%	20%	2%	0%	0%	0%
Improved	0	0	26	29	51	40	30	0	33	0
Unimproved	0	0	7	14	16	40	68	100	67	100

of standard scores with a mean of 50 and a standard deviation unit of 10. In the usual use of the MMPI there is a dual interest. We are interested in distinguishing clinical populations from one another, but we are also interested in selecting abnormals from the general population. For this reason the scoring norms are derived from the general population. On the *Ps* scale, however, interest centers exclusively on clinical groups; hence, the writer departed from MMPI tradition and adopted a form of presentation more in keeping with the nature of the scale. The scale was divided into ten equal parts. Then the percentage of recovered, improved, and unimproved cases falling within each zone of scores was computed. Because the means of the criterion and test groups are so similar (see Tables 2 and 3), the percentages were calculated on the basis of the scores of the 100 test cases and the 84 criterion cases. The gain in stability within each zone of scores

the patients with similar scores recover with shock treatment, 30 per cent improve, and 68 per cent remain unimproved. The onus of whether or not to treat this particular case with shock treatment still rests with the psychiatrist, but he now has a more objective basis for determining the probable success of the treatment. The decision whether or not to shock depends largely on the probability of success with alternative forms of treatment. Until the probabilities for success with various therapies can be determined with equal facility, the *Ps* scale may only serve to heighten the anxieties of a psychiatrist faced with a patient whose prognosis seems hopeless. In the long run, however, the scale can serve more as a stimulant than a deterrent. Knowing that patients refractive to shock treatment can be isolated, workers in the field may be spurred on to devise new forms of treating these patients so that their prognosis can be improved.

### B. APPLICATION OF *Ps* TO A GROUP OF NORMALS

The writer has emphasized that the present scale is intended for use with clinical populations. Therefore, scoring the records of a group of normals on *Ps* has little if anything to do with its validity. *Ps* is not applied to a normal sample, however, merely out of academic curiosity, for this bit of evidence will be used later in discussing theoretical implications of the scale. The MMPI records of 25 medical students and 25 student nurses were scored on *Ps*. Although these samples are admittedly highly selected, they serve the needs of the present study. The mean scores on *Ps* are 15.8 for the males and 16.1 for the females—an average score of 15.9 for the entire group. The small difference in the means of the sexes is not significant. The mean of the normal sample is slightly less than the mean of the recovered test group, which is 17.6. This difference between the means is not significant. Two critical points are arbitrarily set on *Ps* at 22 and 28 points, dividing the scale into three parts labeled recovered, improved, and unimproved. Of the normal sample, 82 per cent fall into the recovered category, 12 per cent into the improved, and 6 per cent into the unimproved. The fact that *Ps* predicts a favorable prognosis for shock therapy for most members of the general population raises theoretical questions about the similarity between recovered patients and people in general. This problem is taken up later.

### C. APPLICATION OF *Ps* TO A GROUP OF PSYCHONEUROTICS

If a scale derived from a group of psychotic patients treated by shock predicts accurately when applied to a group of psychoneurotics treated by psychother-

apy, there is some basis for believing that the factors making patients amenable or refractive to treatment are similar in the two groups. The value of the scale is also enhanced if it is valid for more general purposes. It is true, however, that in psychotherapy more variables operate to affect the results, and in a way that introduces unreliability into the ratings of psychiatric status. Variables like the skill of the therapist and the intricacies of a specific interpersonal relationship are important in determining the outcome of psychotherapy but play relatively little part in shock treatment. Nevertheless, an attempt has been made to collect a group of MMPI records of psychoneurotics treated by brief psychotherapy. For this purpose there were available the records of a group of 53 patients who were the subjects in a special investigation. Four therapists had conducted the treatment of the entire group. The study has been described in detail elsewhere (2) so that the identifying data on the patients will not be repeated. The patients had been given a short form of the MMPI in which all items not appearing on the scales extant at the time were deleted from the test. Since 11 of the deleted items appear on *Ps*, it was necessary to abbreviate *Ps* to 41 items. This procedure precludes any exact comparison between patients subjected to brief psychotherapy and those treated by shock. At the termination of therapy one highly skilled therapist had rank-ordered all 53 patients from the most to the least improved. Making use of these rank orders, the writer arbitrarily divided the cases into three groups, rating the first 18 as responding best to treatment, the next 17 as composing the middle group, and the last 18 as responding most poorly to treatment. These ratings are not quite



TABLE 6  
COMPARISON OF MEANS ON THE *Ps* SCALE FOR PSYCHONEUROTICS  
TREATED BY BRIEF PSYCHOTHERAPY

Groups	<i>N</i>	Mean	<i>SD</i>	<i>Diff</i>	<i>CR</i>	<i>P</i>
I vs. II						
I	18	14.9	6.2	0.1	0.05	.96
II	17	14.8	5.5			
I vs. III						
I	18	14.9	6.2	5.6	2.39	.02
III	18	20.5	7.4			
II vs. III						
II	17	14.8	5.5	5.7	2.54	.01
III	18	20.5	7.4			

comparable to ratings of recovered, improved, and unimproved, for there is no reason to believe that the arbitrary ratings arise in the equal proportions used for the present purpose. The results on the abbreviated form of *Ps* are summarized in Table 6. The difference in scores is not significant between the first two groups, but each of these does differ significantly from the third group at the 2 per cent level of significance. In order to facilitate comparison with the shock groups, the means were weighted to

make them equivalent to full-scale values on *Ps*. For example, in Group I,  $14.9/41 \times 11 = 4.0$ , which makes 18.9 the equivalent score. The other equivalent scores are 18.8 for Group II and 26.0 for Group III. The means of the first two groups are somewhat higher than the mean of the group that recovered with shock treatment. The mean of Group III is between the means of the improved and unimproved shock groups. Although *Ps* does not work as well on the brief psychotherapy groups, it does predict to

TABLE 7  
COMPARISON OF MEANS OF TEST GROUPS ON *Ps* BEFORE  
AND AFTER SHOCK TREATMENT

Group	<i>N</i>	Mean	<i>SD</i>	<i>Diff</i>	<i>CR</i>	<i>P</i>
Recovered						
Preshock	28	18.4	6.0	4.7	2.94	.003
Postshock	28	13.7	6.2			
Improved						
Preshock	26	25.8	5.5	5.8	3.39	.001
Postshock	26	20.0	6.5			
Unimproved						
Preshock	12	29.8	3.5	5.8	2.92	.004
Postshock	12	24.0	5.5			

some extent. There is some evidence that Groups I and II in the brief psychotherapy study (2, p. 274) are quite similar in terms of their MMPI profiles. This similarity in profiles may account in part for their identical scores on *Ps*. Thus, there is a basis for believing that *Ps* might serve as a general prognostic scale for all types of therapy. This sweeping claim would have to be subjected to rigorous study before the scale could be used discriminately. The success of *Ps*, even though limited, on the brief psychotherapy group raises some interesting theoretical questions. The dynamic factors inherent in personality structure which make for good or poor response to therapy would seem to be similar regardless of overt symptomatology. Although this evidence is meager, it is consistent with the results obtained when *Ps* is analyzed on the basis of the type of shock treatment and on the basis of diagnosis.

#### D. APPLICATION OF *Ps* TO POSTSHOCK MMPI RECORDS OF THE TEST GROUPS

It is of some interest to note what happens to scores on *Ps* when the scale is applied to postshock MMPI records. Table 7 presents these results for those members of the test groups who had been given both a pre- and a postshock MMPI. There is a significant drop in the average score on *Ps* for each group from before to after treatment. Thus, it would appear that the norms set up on the basis of pre-shock records cannot be applied to postshock records without incurring gross error. It is probably best not to interpret *Ps* scores from MMPI records taken during the course of treatment, for even a few shock treatments may distort the scores on this scale to a marked degree.

## CHAPTER IV

### ANALYSIS OF THE *Ps* SCALE

#### A. RELIABILITY OF THE *Ps* SCALE

**A**N ODD-EVEN reliability coefficient was computed on the scores of 94 of the test cases.<sup>1</sup> The correlation coefficient obtained from this analysis was .75, which rose to .86 when corrected by the Spearman-Brown formula. The magnitude of this coefficient compares favorably with the reported figures for other MMPI scales.

In all their work on the MMPI, the authors of the test have concentrated on validity rather than reliability. Reliability looms most important, however, in earlier work on most other personality inventories. One can only conjecture why validity has been the area of concentration in MMPI research. Generally high reliability was obtained by using methods of internal consistency to purify the tests, for a homogeneous set of items more readily admits to being measured with a high degree of consistency. There is, however, a serious drawback to using methods of internal consistency to achieve high reliability. The nontest behavior which the test supposedly predicts is not usually of a homogeneous nature. Measures of internal consistency, therefore, often achieve reliability at the expense of validity by eliminating items which predict the criterion but are not consistent with the majority of the items. There is no need to sacrifice validity for reliability, however, for an alternative method of achieving high reliability is to increase the number of items on the test.

<sup>1</sup>The test groups were not quite complete when the reliability coefficient was computed; hence, the omission of six test cases.

Thus, the very items discarded by measures of internal consistency are the ones that might well be retained and, in fact, augmented by similar items, so that each of the subfactors in the scale would be represented in sufficient amount to be measured reliably. If high reliability is a desideratum, it is the procedure of increasing the number of items on the test that should be followed if validity is not to be impaired.

#### B. ITEM OVERLAP AND INTERCORRELATIONS

Although there is a considerable overlap between items on *Ps* and items on all other MMPI scales, items on *Ps* do not comprise a large percentage of any one scale, as shown by Table 8. Of the 52 items, 41 appear on at least one other scale. Most of the overlap occurs on the psychotic scales. Indeed, on the neurotic

TABLE 8  
ITEM OVERLAP BETWEEN *Ps* AND  
OTHER MMPI SCALES

Scale	Number of Items	Com- mon Items	Number Scored Same	Number Scored Opposite
<i>L</i>	15	2	0	2
<i>F</i>	64	3	3	0
<i>K</i>	30	2	0	2
<i>Hs</i>	33	1	0	1
<i>H</i>	55	2	0	2
<i>CH</i>	48	6	6	0
<i>D</i>	60	3	1	2
<i>Hy</i>	60	7	2	5
<i>Pa</i>	50	8	6	2
<i>Pd</i>	62	11	8	3
<i>Mf</i>	60	8	6	2
<i>Pa</i>	40	7	5	2
<i>Pt</i>	48	9	9	0
<i>Sc</i>	78	10	10	0
<i>Ma</i>	46	5	4	1
<i>N</i>	78	9	7	2
Not on other scales	169	11	—	—

scales, the small number of communal items are usually scored in opposite directions. The largest percentage of overlap of *Ps* with any other scale is 18 per cent with *Pt*. It should be noted that 30 per cent of the MMPI items are not found on any of the MMPI scales. Only 21 per cent of the items on *Ps* are also not on another scale; hence, this remaining pool of items on the MMPI contributes a smaller proportion of items to *Ps* than might be expected.

scales also have the highest correlations with *Ps*. Nevertheless, these correlations would still be substantial even if the overlapping items were removed, because, as pointed out previously, the overlap is not great on any one scale. Hence, the generalization can be made that high scores on the psychotic scales tend to go along with a poor prognosis, whereas scores on the neurotic scales seem to have little to do with prognosis. The most surprising correlations are the high negative corre-

TABLE 9  
CORRELATION OF *Ps* WITH THE OTHER SCALES OF THE MMPI ON 100 TEST CASES

<i>L</i>	<i>F</i>	<i>K</i>	<i>Hs</i>	<i>D</i>	<i>Hy</i>	<i>Pd</i>	<i>Mf</i>	<i>Pa</i>	<i>Pt</i>	<i>Sc</i>	<i>Ma</i>	<i>Hcn</i>	<i>Pd</i>
-.42	.54	-.71	.25	-.10	-.12	.52	.11	.34	.62	.72	.59	-.26	.57

The next step was to compute the correlations between *Ps* and the other scales on the MMPI. The records of the 100 test cases were used because they represent the population on whom the scale is to be used. The results are given in Table 9. The intercorrelations are given for the published scales and for several scales which are no longer in general use. A factor which tends to lower the correlation coefficients is the marked skewness of most of the distributions, the Pearsonian  $r$  not being a maximally descriptive statistic in such a case. A correlation coefficient with a magnitude of .25 is significant at the 1 per cent level with a sample of this size.

The *D*, *Hy*, and *Mf* scales fail to meet this level of significance; in fact, all three are far short of being significant. The other scales, however, have a highly significant correlation with *Ps*. Although skewness tends to lower the intercorrelations, the item overlap between scales may spuriously raise them. Common elements are greatest on the psychotic and the psychopathic deviate scales, which

lations of *Ps* with both *L* and *K*, which cannot be considered to be spuriously high since the item overlap is negligible with both scales. The *K* scale has been interpreted by Meehl and Hathaway (10) as a correction for an unconscious tendency of the respondent to put himself in either a favorable or unfavorable light. A high *K* score calls for an upward revision of several of the MMPI scales. In the present study, however, a high *K* score goes along with a favorable prognosis. This fact seems to cast some doubt on the generality of *K*, and on the validity of the authors' recommendation that all records be systematically corrected for *K*. From the evidence in the present study, it seems that *K* has a different meaning for these patients selected for shock treatment—in fact, just the opposite meaning from that intended by the authors. It may be that *K* works best on the subgroup for whom it was originally designed, namely, clinical populations with normal profiles; hence, the *K* correction might best be restricted to this group.



TABLE 10  
MEANS AND STANDARD DEVIATIONS ON *Ps* OF TEST GROUPS IN RELATION  
TO THE TYPE OF SHOCK TREATMENT

Treatment	Recovered			Improved			Unimproved		
	<i>N</i>	Mean	<i>SD</i>	<i>N</i>	Mean	<i>SD</i>	<i>N</i>	Mean	<i>SD</i>
Electric shock	25	17.6	6.4	22	23.3	5.5	7	28.4	4.5
Electronarcosis	6	18.3	6.3	5	25.0	3.2	4	27.8	5.9
Insulin	7	17.7	6.5	13	26.9	5.2	8	31.0	5.5

### C. RELATION OF *Ps* TO TYPE OF SHOCK

In the derivation of *Ps*, patients were lumped together without regard for the type of treatment received. Patients in the study were treated with insulin, electronarcosis, or electric shock. An objection might be raised that the use of mixed groups obscures results inasmuch as the efficiency of the treatment is supposed to vary with the type of patient. In this study the average score on *Ps* for all patients treated with electric shock is 21.3, and for all patients treated with insulin, 25.8. This difference is statistically significant. The average score on *Ps* for all patients treated with electronarcosis is 23.1, which does not differ significantly from that of either of the other two groups. If it is true, however, that insulin produces better results on patients with a poor prognosis than does electrical treatment, then patients treated with insulin should have more favorable ratings of psychiatric status with correspondingly higher scores on *Ps*, and the reverse should hold for patients treated electrically. Hence, a se-

ries of tests was made to determine the effect of the type of treatment. All 40 test cases rated recovered clinically were divided into three groups on the basis of the type of treatment. The groups were then compared with one another by means of the *t*-test (11). The improved and unimproved test groups were similarly analyzed. The results are collected in Tables 10 and 11. None of the nine comparisons reveals a significant difference. The type of treatment does not seem to affect predictions on *Ps*. These results based upon so few cases are certainly not conclusive, although they are highly suggestive. These results may, however, be affected to an unknown degree by the clinical judgments that went into deciding why patients were to receive the treatment which was actually administered. Nevertheless, the results lend credence to the notion that all three treatments are equally efficacious. To obtain more unequivocal results, an equal number of matched patients would have to be given the three treatments and the choice of treatment would have to be

TABLE 11  
TESTS FOR SIGNIFICANCE BETWEEN TEST GROUPS RATED THE SAME CLINICALLY BUT  
TREATED WITH DIFFERENT KINDS OF SHOCK

Treatment	Recovered			Improved			Unimproved		
	<i>df</i>	<i>t</i>	<i>P</i>	<i>df</i>	<i>t</i>	<i>P</i>	<i>df</i>	<i>t</i>	<i>P</i>
ECT vs. I	30	0.04	.90	33	1.85	.065	15	1.13	.20
ECT vs. EN	29	0.24	.90	25	0.63	.50	9	0.17	.80
EN vs. I	11	0.15	.80	16	0.73	.40	10	1.03	.30

TABLE 12  
MEANS AND STANDARD DEVIATIONS ON  $P_s$  OF TEST GROUPS  
IN RELATION TO DIAGNOSIS

Treatment	Recovered			Improved			Unimproved		
	<i>N</i>	Mean	<i>SD</i>	<i>N</i>	Mean	<i>SD</i>	<i>N</i>	Mean	<i>SD</i>
Aff. dis.	23	16.4	5.9	15	23.2	4.4	1	30.0	—
Schiz.	13	20.1	5.2	17	26.4	5.9	12	30.4	4.0
Psychoneur.	2	18.5	5.5	8	23.9	4.7	6	27.2	5.9

made solely on the basis of the matchings. Incidentally, the effect of psychotherapy on psychotic patients could be assayed in the same way. It might also be pointed out that  $P_s$  furnishes a rational basis for selecting equated groups to put this suggested hypothesis to a more rigorous test.

#### D. RELATION OF $P_s$ TO DIAGNOSIS

As stated earlier, diagnosis was also ignored in selecting prognostic groups.

the unimproved group includes more schizophrenics. The average score on  $P_s$  is 19.4 for all patients diagnosed as affective disorders, 24.4 for all psychoneurotics, and 25.6 for all schizophrenics. The differences between the average scores of the affective disorders and of each of the other two groups are significant, but the difference between the schizophrenic and the psychoneurotic groups is not significant. Nevertheless, if diagnosis as such is related to test results,

TABLE 13  
TESTS FOR SIGNIFICANCE BETWEEN DIFFERENT DIAGNOSTIC GROUPS  
WITHIN THE SAME PROGNOSTIC CATEGORY

Diagnosis	Recovered			Improved			Unimproved		
	<i>df</i>	<i>t</i>	<i>P</i>	<i>df</i>	<i>t</i>	<i>P</i>	<i>df</i>	<i>t</i>	<i>P</i>
Aff. vs. Sc	34	1.82	.07	30	1.68	.11	11	0.09	.90
Aff. vs. PN	23	0.47	.60	21	0.34	.70	5	0.40	.70
PN vs. Sc	13	0.37	.70	23	1.00	.30	6	1.28	.20

An objection might be raised that  $P_s$  merely discriminates the affective disorders from schizophrenics. This proposition can be subjected to the same type of test as was done for the effect of the type of treatment. Each prognostic group contains cases of affective disorders,<sup>2</sup> schizophrenics, and severe psychoneurotics. To be sure, the recovered group includes more affective disorders whereas

scores on  $P_s$  within a single prognostic category should differ significantly when subgroups are compared on the basis of diagnosis.

The records of the 100 test cases were again utilized, but three cases were omitted because they were not assigned a definite diagnosis. The means of all subgroups were analyzed by *t*-tests to determine whether any of the differences achieves statistical significance. The results are presented in Tables 12 and 13. None of the differences is statistically significant; therefore, differential diagnosis seems to have little effect on the predic-

<sup>2</sup> The affective disorders include manic-depressive and involuntional patients because the latter are too few in number to be analyzed separately. Most of the patients in the psychoneurotic category were diagnosed as obsessive-compulsive or severe anxiety hysteria.



tive power of *Ps*. It should be pointed out that some of the subgroups—especially the unimproved affective disorders—are composed of very few cases. This fact decreases the exactness of the comparisons. The scores of the schizophrenic group are consistently, even though not significantly, higher than those of other groups. Nevertheless, in terms of the arbi-

trary threefold classification of improvement on *Ps*, the outcome for the majority of schizophrenics is correctly predicted. Eight of the 13 recovered, 10 of the 17 improved, and 9 of the 12 unimproved schizophrenics are placed correctly by *Ps*. Of the entire group of 42 schizophrenic test cases, only 1 is misplaced by two steps in the three-step classification on *Ps*.

## CHAPTER V

### DISCUSSION AND INTERPRETATION OF *Ps*

ONE OF the aims of the present study is to gain insight into the differences between the prognostic groups. Any a posteriori explanation of data is, of course, subject to question. The real test of explanations is whether or not they lead to further formulations which can be subsequently studied.

The item content furnishes many clues about the nature of the prognostic groups. In the first place, *Ps* contains an *unusually large number of items which refer to interpersonal relationships* and only a few items which refer to symptomatic complaints, although the entire pool of items on the MMPI contains almost equal numbers of each kind. That is, the items on *Ps* refer to interpersonal relations in the family, at work, and socially, rather than to physical ailments, fears, mood disorders, or disorders of thought. In so far as overt behavior can be inferred from responses to items, it can be conjectured that the presenting symptomatic picture of members of the various prognostic groups is quite similar, and that the chief differences between the groups will be found in the dynamics of genetic development and personality structure.

Carrying this point somewhat further, patients who subsequently recover respond to these particular items even before treatment in much the same way as a group of normals. This is shown by the fact that the recovered and normal groups have almost identical average scores on *Ps*. The mean of the recovered group is 17.6, and the mean of the selected sample of normals is 15.9. A comparison of the item counts of the recovered group with the original normative

data of the MMPI<sup>1</sup> reveals that only 14 of the 52 items differ by as much as 20 per cent, which is approximately at the 5 per cent level throughout most of the range of proportions. On 9 of these 14 items, the item counts of the normal group are actually higher in the "plus" direction. In contrast, 38 of the 52 items differ between the unimproved and the normal group, and in all but 5 of these, the unimproved group has the higher percentage of "plus" responses. Each member of the prognostic groups must have at least one elevated score on the other MMPI scales; yet on this set of prognostic items, the responses of the recovered group are similar to those of the normals. This finding becomes elucidated, in part, when some theoretical sense is made of the content of the items.

The answers in the unimproved direction on *Ps* are consistent with the picture of a person who has withdrawn from a world perceived as hostile. These respondents express open dissatisfaction with the family group, e.g., "I believe that my home life is as pleasant as that of most people I know" (F), and "At times I have very much wanted to leave home" (T). It can be conjectured that the poor identifications established in the family group have been generalized to other authority figures as shown by, "I am apt to pass up something that I want to do because others feel that I am not going about it in the right way" (T), and "I have often had to take orders from someone who did not know as much as I did" (T). Unimproved patients seem to have very poor object relations. They

<sup>1</sup>The writer is indebted to the authors of the MMPI for making these data available.

perceive other people as being exploitative, unfriendly, and hostile, e.g., "Most people make friends because friends are likely to be useful to them" (T), and "People say insulting things about me" (T). There seems to be nothing about their lives that enhances their self-esteem; consequently, they express marked feelings of inadequacy, e.g., "In a group of people I would not be embarrassed to be called upon to start a discussion or give an opinion about something I know well" (F), and "I am not unusually self-conscious" (F).

These people seem to have no sources of ego strength. They feel unable to cope with their environments. Also, they feel dependent upon arbitrary external forces which show no respect for their rights. This theme is summed up by the statements, "People can pretty easily change me even though I thought my mind was made up on a subject," and "No one cares much what happens to you," which are answered "True" by patients remaining unimproved. They seem to recognize their sorry plight (they can see that everything is wrong, including themselves); but since they are unable to cope with their problems, they have no motivation to get well. It seems quite possible that many of them feel that a psychotic adjustment is the best possible form of adjustment, for at least they are relatively free from the demands of a hostile environment. It is interesting to note that as a group they tend to have low *K* scores. Hathaway (3) has suggested that *K* can be thought of as a continuum from capacity for reflective self-analysis at the low end to defensiveness and obtuseness at the high end. The only possible way that insight can be attributed to the unimproved group is in the trivial sense that they realize that in view of their unfortunate circumstances, there is no point

in trying to get well. This, of course, is not the usual connotation of insight.

The patients in the recovered group, on the other hand, are still trying to establish good object relations. In view of their psychotic condition there is good reason to doubt the verity of many of their statements. The important thing, however, is that they perceive their environment as being relatively friendly and rewarding, and they feel that they are in better rapport with their family, friends, and self. Their attitude with respect to their interpersonal relations seems Pollyannish; hence, their high *K* scores. They can admit their symptoms but must repress their underlying interpersonal conflicts.

By definition, individuals with severe personality disorders have poor interpersonal relations. Although the answers on *Ps* of the recovered group are similar to the answers of normals, it should not be concluded that the recovered group are similar to the normals in ego strength or in the capacity to form good object relations. Rather, the answers of the recovered group might be thought of as representing a potentiality for greater ego strength and better object relations. This potentiality seems to be lacking in the unimproved group. However, the responses of the recovered patients may reflect the actuality of greater ego strength and better object relations, even before treatment, than holds in the case of the unimproved patients. At any rate, *patients who recover perceive that there are many worthwhile things in life for them; so they have motivation to recover. Their motivation is also enhanced by the fact that they perceive themselves as being capable of coping with many of their problems.*

It should be apparent that the preceding analysis is of the momentary dynam-

ics of the two groups. We should like to know the particular family constellations that led to the different motivational systems. What causes the unimproved group to have such vulnerable egos? What are the sources of ego strength that give some patients a greater propensity to recover from a psychotic condition? Unfortunately, these questions cannot be answered from an analysis of the properties of the present scale. The answers to such questions must come from long-range genetic studies and intensive interview material.

The differences in response to items on *Ps* may also cast some light on the nature of current nosological systems. The present results indicate that there are wide differences in personality structure among a group of individuals with highly similar presenting symptoms. The correlations of *Ps* with other MMPI scales (Table 9) indicate that unimproved patients are characterized by high scores on the psychotic and psychopathic deviate scales, whereas recovered patients have significantly lower scores on these scales; yet both groups are composed chiefly of psychotics. Evidently, patients who subsequently recover are diagnosed as psychotics clinically on the basis of their overt symptoms, and little attention is paid to their interpersonal relations. It would seem to follow that there will be many discrepancies between diagnoses based upon the MMPI and those based upon clinical judgment. However, many of the discrepancies would arise from the nature of the diagnostic categories in current use. Although these discrepancies limit the efficiency of the test as a diagnostic instrument, the limitations may be more apparent than real, for it may be that what is needed is a modification of the present system of diagnosis.

It is seen from Table 11 that *Ps* cuts across diagnostic groups rather effectively. This is not the first instance of a prognostic test doing so. The Elgin Prognostic Scale for Schizophrenia (15) works equally well in predicting the outcome of manic-depressives. Then, too, *Ps* has some success in separating the brief psychotherapy group. There is a growing dissatisfaction with the Kraepelinian nosology for many reasons. Chief of these is the fact that so many cases present such diffuse symptoms that they defy clear-cut diagnosis. If we examine the reasons why it is desirable to classify subgroups of psychiatric patients, the immediate purpose seems to be to

convey a lot of information in a few summary terms. This implies that a body of knowledge is correlated with each of the divisions made. If the divisions are equivocal, the probability of relating them meaningfully to other variables is reduced. The proper classificatory system, then, is the one that summarizes the most knowledge by virtue of its high correlation with a large body of meaningful facts. Not only should the categories be correlated with known facts, but they should be convenient divisions for further study.

The present psychiatric classifications are phenotypic in that diagnosis is based largely on overt symptomatology, i.e., physical complaints, fears, mood swings, thought content, etc. The Kraepelinian system undoubtedly has some merit, as witnessed by the tenacity with which this nosology is retained. It may be, however, that psychotic patients who recover with treatment form a more genotypic group than do patients classified into one of the diagnostic categories, e.g., schizophrenia. By this the writer means that a classificatory system based upon prognosis rather than current diagnostic categories may be collated with other variables in a more meaningful way than the present system allows.

Essential agreement with this position is shared by Wittman (16), who distinguishes "process" schizophrenia from "schizophreniform" on the basis of prognosis. The latter embraces cases whose symptomatology resembles schizophrenia but who have a good prognosis because they do not have the stormy childhood and difficult adolescence which tend to lead to the withdrawal and indifference characteristic of the "process" schizophrenic. She compared 66 adult schizophrenics who as children had been referred to a clinic for study, with a similar number of schizophrenics who had no known history of childhood referral. The groups differed markedly on the Elgin Prognosis Scale; the former had a much poorer prognosis. Wittman cites this study as one bit of evidence for believing that it is fruitful to divide schizophrenics into the two groups she suggests. The study also points up the fact that thinking in terms of prognosis rather than diagnosis leads to productive research.

There is a great deal of equivocal data in the field of diagnostic testing which would probably be elucidated in part if the data were analyzed on the basis of prognosis. The writer does not believe that it is either advisable or feasible to dispense with the present system of nosology. He does believe that a dual system of classification could be employed in most psychological studies whereby results would be analyzed in terms of prognosis as well as diagnosis. By this means, some of the defects of present diagnostic groupings might be circumvented.



## CHAPTER VI

### THE LOW-PROFILE GROUP

IN STANDARDIZING *Ps* certain requirements were established for including cases in the criterion groups. About 20 per cent of the cases failed to meet the restriction placed on the magnitude of the *T*-scores on the personality variables of the MMPI. These excluded cases had *T*-scores below 70 on all scales or on all scales except *Pd*, or *Ma*, but not both. This excluded group numbers 44 in all. The results on this group are now presented.

When a breakdown for clinical ratings was made, 20 were rated recovered, 20 unimproved, and only 4 were rated improved. The rating procedure was exactly the same as in the main study. The groups were somewhat younger on the average than the main groups; the recovered group had a mean age of 30.3 years, whereas the unimproved group had a mean of 25.6. The groups were also a little more intelligent than the main groups; the average IQ on the short form of the Wechsler-Bellevue was 118 for the recovered and 108 for the unimproved. There was a preponderance of females in both groups, 13 in the recovered group and 19 in the unimproved group. The proportion of females in the unimproved group was much higher, whereas the sex ratio in the recovered group was quite similar to that in the main study. The distributions of diagnoses were most dissimilar to those in the main study. The recovered group was split evenly, 10 schizophrenics and 10 affective disorders, whereas the unimproved group contained 18 schizophrenics and 1 each of affective disorders and psychoneurotics. It is difficult to evaluate the meaning of

these personal factors, but it is interesting to note that a great many of the schizophrenics were of the paranoid type.

#### A. SCORES ON *Ps*

The records of the recovered and unimproved groups were first scored on *Ps*. The average score of the recovered group was 14.1 and of the unimproved group, 13.3. These scores do not differ significantly. Both means are actually lower than the mean of the recovered group in the main study. Furthermore, 37 of the 40 cases would be predicted to recover by *Ps*. Thus, it is apparent that *Ps* has little value in predicting the results of this selected group.

#### B. ITEM ANALYSIS

In spite of the small-sized samples and the lack of an independent test group, an item analysis was carried out between the records of the 20 recovered and the 20 unimproved cases. Only 17 items differed significantly between the two groups at the 5 per cent level. This number is even less than would be expected solely through errors of random sampling. Only a few of the items were significant at the 2 per cent level. With such barren results, there was no reasonable chance that these items comprised into a scale would predict the outcome of treatment better than chance for the groups in question. Therefore, the attempt to predict the outcome of shock treatment for this subgroup was abandoned, at least for the time being.

The writer does not believe that these cases are merely test "misses" in the sense that their psychosis should have

been reflected in high scores on the existing scales of the MMPI. Their ability to respond to the test in a way that avoids high T-scores probably reflects essential differences in personality structure. Cer-

tainly, this group should be challenging to many of our notions about psychopathology, especially if it is found that they simulate normality on other personality measures.



## CHAPTER VII

### SUMMARY AND CONCLUSIONS

1. The aim of this research was to devise a scale to predict the probability of success of shock treatment. This was done by the empirical selection of MMPI items on which the responses of a criterion group of recovered patients differed significantly from the responses of a criterion group of unimproved patients.

2. The resulting scale of 52 items selected by the item analysis was labeled *Ps*. This scale not only discriminated the original criterion groups with a high degree of accuracy but worked almost as well on independent test groups. The scale predicted accurately enough to warrant its use for individual prediction.

3. The magnitude of the odd-even reliability coefficient was fairly high, comparing favorably with the reliability of other MMPI scales.

4. The scale significantly differentiated a group of psychoneurotics with respect to outcome of treatment by brief psychotherapy, but the overlap between the outcome groups was too great for this finding to be anything but suggestive.

5. The correlation of scores on *Ps* with other MMPI scales varied considerably. It had uniformly high positive correlations with the psychotic scales, little relationship to the neurotic scales, and high negative correlations with *L* and *K*. These latter correlations called into question the general applicability of the *K* scale.

6. Scores on *Ps* do not seem to be af-

fected materially by either the type of shock treatment or the diagnosis. The inherent propensity to improve seems to be the main variable affecting scores on *Ps*.

7. Members of the unimproved group were characterized as having little ego strength and little motivation to improve, whereas members of the recovered group were characterized as persons still struggling to maintain object relations as well as having greater motivation to improve.

8. A separate analysis was made of a subgroup of patients with essentially normal MMPI profiles. No effective way was found of predicting their response to treatment.

The *Ps* scale is certainly not the last word in predicting response to therapy. Nevertheless, it works well enough to furnish a practical tool for use with a fairly wide range of patients. What is even more important is that the scale may be used for other research problems:

a) It may serve as a rational basis for selecting matched groups of patients in comparing the effectiveness of various types of treatment, e.g., insulin vs. electric shock.

b) It may be used to isolate patients with a poor prognosis, for the purpose of finding better techniques of treatment.

c) It may also clarify some aspects of current psychiatric nosology by pointing out basic similarities among groups differently classified by the present system.

# APPENDIX DATA ON PROGNOSTIC GROUPS

## A. HOSPITALIZATION

Group	N*	Per Cent Previously Treated by Shock	Per Cent Subsequently Hospitalized
Recovered criterion	42	9	21
Unimproved criterion	42	14	62
Recovered test	40	18	18
Improved test	40	25	10
Unimproved test	20	10	15
Low-profile recovered	20	20	15
Low-profile unimproved	20	20	45

\* The number of cases in each group in B, C, and D is the same as in the groups in A.

## B. PERSONAL DATA

Group	Average Age	Sex		Preshock Intelligence	
		M	F	N	Mean
Recovered criterion	37.5	31%	69%	33	109
Unimproved criterion	32.0	52	48	26	103
Recovered test	38.5	43	57	15	114
Improved test	34.5	28	72	14	107
Unimproved test	30.5	40	60	4	102
Low-profile recovered	30.3	35	65	9	118
Low-profile unimproved	25.6	5	95	10	108

## C. DIAGNOSIS

Group	Affective Disorder	Schizo- phrenic	Psycho- neurotic	Un- known
Recovered criterion	69%	26%	5%	0%
Unimproved criterion	22	52	26	0
Recovered test	58	31	5	5
Improved test	38	42	20	0
Unimproved test	5	60	30	5
Low-profile recovered	50	50	0	0
Low-profile unimproved	5	90	5	0

## D. TYPE OF SHOCK TREATMENT

Group	Electric Shock	Electro- narcosis	Insulin	Un- known
Recovered criterion	85%	2%	13%	0%
Unimproved criterion	43	17	40	0
Recovered test	62	15	18	5
Improved test	55	12	33	0
Unimproved test	35	20	40	5
Low-profile recovered	75	5	20	0
Low-profile unimproved	20	15	65	0

## REFERENCES

1. EDGERTON, H. E., AND PATERSON, D. G. Table of standard errors and probable errors of percentages for varying numbers of cases. *J. appl. Psychol.*, 1926, 10, 378-391.
2. HARRIS, R. E., AND CHRISTIANSEN, CAROLE. Prediction of response to brief psychotherapy. *J. Psychol.*, 1946, 21, 269-284.
3. HATHAWAY, S. R. Some considerations relative to nondirective counseling, as therapy. *J. clin. Psychol.*, 1948, 4, 226-231.
4. HATHAWAY, S. R., AND MCKINLEY, J. C. A multiphasic personality schedule: I. Construction of the schedule. *J. Psychol.*, 1940, 10, 249-254.
5. HATHAWAY, S. R., AND MCKINLEY, J. C. A multiphasic personality schedule: III. The measurement of symptomatic depression. *J. Psychol.*, 1942, 14, 73-74.
6. HATHAWAY, S. R., AND MCKINLEY, J. C. *Manual for the Minnesota Multiphasic Personality Inventory*. Minneapolis: Univ. Minn. Press, 1943.
7. MCKINLEY, J. C., AND HATHAWAY, S. R. A multiphasic personality schedule: II. A differential study of hypochondriasis. *J. Psychol.*, 1940, 10, 255-268.
8. MCKINLEY, J. C., AND HATHAWAY, S. R. A multiphasic personality schedule: IV. Psychasthenia. *J. appl. Psychol.*, 1942, 26, 614-624.
9. MCKINLEY, J. C., AND HATHAWAY, S. R. The Minnesota Multiphasic Personality Inventory: V. Hysteria, hypomania, and psychopathic deviate. *J. appl. Psychol.*, 1944, 24, 153-174.
10. MEEHL, P. E., AND HATHAWAY, S. R. The K factor as a suppressor variable in the Minnesota Multiphasic Personality Inventory. *J. appl. Psychol.*, 1946, 30, 525-564.
11. RIDER, P. R. *An introduction to modern statistical methods*. New York: Wiley, 1949.
12. SCHNECK, J. M. Clinical evaluation of the F scale on the MMPI. *Amer. J. Psychiat.*, 1948, 104, 440-442.
13. STRONG, E. K. *Vocational interests of men and women*. Palo Alto, Calif.: Stanford Univ. Press, 1942.
14. TRELOAR, A. *Elements of statistical reasoning*. New York: Wiley, 1939.
15. WITTMAN, M. PHYLLIS. A scale for measuring prognosis in schizophrenic patients. *Elgin Pap.*, 1941, 4, 20-33.
16. WITTMAN, M. PHYLLIS. Diagnostic and prognostic significance of the shut-in personality type as a prodromal factor in schizophrenia. *J. clin. Psychol.*, 1948, 4, 211-214.